

SEQUENCE LISTING

(1) GENERAL INFORMATION:

- 5 (i) APPLICANT: Godowski, Paul J.
Mark, Melanie R.
Scadden, David T.
Baker, Kevin P.
Baron, Will F.
- 10 (ii) TITLE OF INVENTION: Protein Tyrosine Kinases
- (iii) NUMBER OF SEQUENCES: 35
- 15 (iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: Genentech, Inc.
(B) STREET: 460 Point San Bruno Blvd
(C) CITY: South San Francisco
20 (D) STATE: California
(E) COUNTRY: USA
(F) ZIP: 94080
- 25 (v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: patin (Genentech)
- 30 (vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER:
(B) FILING DATE: 20-DEC-1993
(C) CLASSIFICATION:
- 35 (vii) PRIOR APPLICATION DATA:
(A) APPLICATION NUMBER: 08/157563
(B) FILING DATE: 23-NOV-1993
- 40 (viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: Hasak, Janet E.
(B) REGISTRATION NUMBER: 28,616
(C) REFERENCE/DOCKET NUMBER: 854C1
- 45 (ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 415/225-1896
(B) TELEFAX: 415/952-9881
(C) TELEX: 910/371-7168

(2) INFORMATION FOR SEQ ID NO:1:

- 50 (i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3611 bases
(B) TYPE: nucleic acid

(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

5

CCGCCGATGG CGCTGAGGCG GAGCATGGGG CGGCCGGGGC TCCCGCCGCT 50

10

GCCGCTGCCG CCGCCACCGC GGCTCGGGCT GCTGCTGGCG GCTCTGGCTT 100

CTCTGCTGCT CCCGGAGTCC GCCGCCGCAG GTCTGAAGCT CATGGGAGCC 150

15

CCGGTGAAGC TGACAGTGTC TCAGGGGCAG CCGGTGAAGC TCAACTGCAG 200

TGTGGAGGGG ATGGAGGAGC CTGACATCCA GTGGGTGAAG GATGGGGCTG 250

TGGTCCAGAA CTTGGACCAG TTGTACATCC CAGTCAGCGA GCAGCACTGG 300

ATCGGCTTCC TCAGCCTGAA GTCAGTGGAG CGCTCTGACG CCGGCCGGTA 350

CTGGTGCCAG GTGGAGGATG GGGGTGAAAC CGAGATCTCC CAGCCAGTGT 400

GGCTCACGGT AGAAGGTGTG CCATTTTTC AAGTGGAGCC AAAAGATCTG 450

GCAGTGCCAC CCAATGCCCC TTTCCAACCTG TCTTGTGAGG CTGTGGGTCC 500

CCCTGAACCT GTTACCATTG TCTGGTGGAG AGGAACTACG AAGATCGGGG 550

GACCCGCTCC CTCTCCATCT GTTTTAAATG TAACAGGGGT GACCCAGAGC 600

ACCATGTTTT CCTGTGAAGC TCACAACCTA AAAGGCCTGG CCTCTTCTCG 650

CACAGCCACT GTTACCTTC AAGCACTGCC TGCAGCCCCC TTCAACATCA 700

CCGTGACAAA GCTTTCCAGC AGCAACGCTA GTGTGGCCTG GATGCCAGGT 750

GCTGATGGCC GAGCTCTGCT ACAGTCCTGT ACAGTTCAGG TGACACAGGC 800

0123456789-0123456789

CCCAGGAGGC TGGGAAGTCC TGGCTGTTGT GGTCCCTGTG CCCCCCTTTA 850

5 CCTGCCTGCT CCGGGACCTG GTGCCTGCCA CCAACTACAG CCTCAGGGTG 900

CGCTGTGCCA ATGCCTTGGG GCCCTCTCCC TATGCTGACT GGGTGCCCTT 950

10 TCAGACCAAG GGTCTAGCCC CAGCCAGCGC TCCCCAAAAC CTCCATGCCA 1000

TCCGCACAGA TTCAGGCCTC ATCTTGGAGT GGAAGAAGT GATCCCCGAG 1050

15 GCCCCTTTGG AAGGCCCCCT GGGACCCTAC AAAGTGCCT GGGTTCAAGA 1100

20 CAATGGAACC CAGGATGAGC TGACAGTGGG GGGGACCAGG GCCAATTTGA 1150

CAGGCTGGGA TCCCCAAAAG GACCTGATCG TACGTGTGTG CGTCTCCAAT 1200

25 GCAGTTGGCT GTGGACCCTG GAGTCAGCCA CTGGTGGTCT CTTCTCATGA 1250

CCGTGCAGGC CAGCAGGGCC CTCCTCACAG CCGCACATCC TGGGTACCTG 1300

TGGTCCTTGG TGTGCTAACG GCCCTGGTGA CGGCTGCTGC CCTGGCCCTC 1350

30 ATCCTGCTTC GAAAGAGACG GAAAGAGACG CGGTTTGGGC AAGCCTTTGA 1400

CAGTGTCATG GCCCGGGGAG AGCCAGCCGT TCACTTCCGG GCAGCCCGGT 1450

40 CCTTCAATCG AGAAAGGCCC GAGCGCATCG AGGCCACATT GGACAGCTTG 1500

GGCATCAGCG ATGAACTAAA GGAAAACTG GAGGATGTGC TCATCCCAGA 1550

45 GCAGCAGTTC ACCCTGGGCC GGATGTTGGG CAAAGGAGAG TTTGGTTCAG 1600

50 TGCAGGAGGC CCAGCTGAAG CAAGAGGATG GCTCCTTTGT GAAAGTGGCT 1650

TACATACTCA CCCCCGGAGG GCTGGCTGAG CAGCCAGGGC AGGCAGAGCA 2600
 5 CCAGCCAGAG AGTCCCCTCA ATGAGACACA GAGGCTTTTG CTGCTGCAGC 2650
 AAGGGCTACT GCCACACAGT AGCTGTTAGC CCACAGGCAG AGGGCATCGG 2700
 10 GGCCATTTGG CCGGCTCTGG TGGCCACTGA GCTGGCTGAC TAAGCCCCGT 2750
 CTGACCCCAG CCCAGACAGC AAGGTGTGGA GGCTCCTGTG GTAGTCCTCC 2800
 15 CAAGCTGTGC TGGGAAGCCC GGA CTGACCA AATCACCCAA TCCCAGTTCT 2850
 TCCTGCAACC ACTCTGTGGC CAGCCTGGCA TCAGTTTAGG CCTTGGCTTG 2900
 ATGGAAGTGG GCCAGTCCTG GTTGTCTGAA CCCAGGCAGC TGGCAGGAGT 2950
 25 GGGGTGGTTA TGTTTCCATG GTTACCATGG GTGTGGATGG CAGTGTGGGG 3000
 AGGGCAGGTC CAGCTCTGTG GGCCCTACCC TCCTGCTGAG CTGCCCCTGC 3050
 TGCTTAAGTG CATGCATTGA GCTGCCTCCA GCCTGGTGGC CCAGCTATTA 3100
 30 CCACACTTGG GGTTTAAATA TCCAGGTGTG CCCCTCCAAG TCACAAAGAG 3150
 ATGTCCTTGT AATATTCCCT TTTAGGTGAG GGTGTAAG GGGTTGGTAT 3200
 40 CTCAGGTCTG AATCTTCACC ATCTTTCTGA TTCCGCACCC TGCCTACGCC 3250
 AGGAGAAGTT GAGGGGAGCA TGCTTCCCTG CAGCTGACCG GGTCACACAA 3300
 45 AGGCATGCTG GAGTACCCAG CCTATCAGGT GCCCTCTTC CAAAGGCAGC 3350
 50 GTGCCGAGCC AGCAAGAGGA AGGGGTGCTG TGAGGCTTGC CCAGGAGCAA 3400

GTGAGGCCCG AGAGGAGTTC AGGAACCCTT CTCCATACCC ACAATCTGAG 3450

CACGCTACCA AATCTCAAAA TATCCTAAGA CTAACAAAGG CAGCTGTGTC 3500

TGAGCCCAAC CCTTCTAAAC GGTGACCTTT AGTGCCAACT TCCCCTCTAA 3550

CTGGACAGCC TCTTCTGTCC CAAGTCTCCA GAGAGAAATC AGGCCTGATG 3600

AGGGGGAATT C 3611

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 890 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Met Ala Leu Arg Arg Ser Met Gly Arg Pro Gly Leu Pro Pro Leu
1 5 10 15

Pro Leu Pro Pro Pro Pro Arg Leu Gly Leu Leu Leu Ala Ala Leu
20 25 30

Ala Ser Leu Leu Leu Pro Glu Ser Ala Ala Ala Gly Leu Lys Leu
35 40 45

Met Gly Ala Pro Val Lys Leu Thr Val Ser Gln Gly Gln Pro Val
50 55 60

Lys Leu Asn Cys Ser Val Glu Gly Met Glu Glu Pro Asp Ile Gln
65 70 75

Trp Val Lys Asp Gly Ala Val Val Gln Asn Leu Asp Gln Leu Tyr
80 85 90

Ile Pro Val Ser Glu Gln His Trp Ile Gly Phe Leu Ser Leu Lys
95 100 105

Ser Val Glu Arg Ser Asp Ala Gly Arg Tyr Trp Cys Gln Val Glu
110 115 120

Asp Gly Gly Glu Thr Glu Ile Ser Gln Pro Val Trp Leu Thr Val
125 130 135

003369.01599
66950

	Glu Gly Val Pro Phe Phe Thr Val Glu Pro Lys Asp Leu Ala Val	140	145	150
5	Pro Pro Asn Ala Pro Phe Gln Leu Ser Cys Glu Ala Val Gly Pro	155	160	165
	Pro Glu Pro Val Thr Ile Val Trp Trp Arg Gly Thr Thr Lys Ile	170	175	180
10	Gly Gly Pro Ala Pro Ser Pro Ser Val Leu Asn Val Thr Gly Val	185	190	195
	Thr Gln Ser Thr Met Phe Ser Cys Glu Ala His Asn Leu Lys Gly	200	205	210
15	Leu Ala Ser Ser Arg Thr Ala Thr Val His Leu Gln Ala Leu Pro	215	220	225
	Ala Ala Pro Phe Asn Ile Thr Val Thr Lys Leu Ser Ser Ser Asn	230	235	240
	Ala Ser Val Ala Trp Met Pro Gly Ala Asp Gly Arg Ala Leu Leu	245	250	255
25	Gln Ser Cys Thr Val Gln Val Thr Gln Ala Pro Gly Gly Trp Glu	260	265	270
	Val Leu Ala Val Val Val Pro Val Pro Pro Phe Thr Cys Leu Leu	275	280	285
30	Arg Asp Leu Val Pro Ala Thr Asn Tyr Ser Leu Arg Val Arg Cys	290	295	300
	Ala Asn Ala Leu Gly Pro Ser Pro Tyr Ala Asp Trp Val Pro Phe	305	310	315
	Gln Thr Lys Gly Leu Ala Pro Ala Ser Ala Pro Gln Asn Leu His	320	325	330
40	Ala Ile Arg Thr Asp Ser Gly Leu Ile Leu Glu Trp Glu Glu Val	335	340	345
	Ile Pro Glu Ala Pro Leu Glu Gly Pro Leu Gly Pro Tyr Lys Leu	350	355	360
45	Ser Trp Val Gln Asp Asn Gly Thr Gln Asp Glu Leu Thr Val Glu	365	370	375
	Gly Thr Arg Ala Asn Leu Thr Gly Trp Asp Pro Gln Lys Asp Leu	380	385	390

Residue	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
	Ile	Val	Arg	Val	Cys	Val	Ser	Asn	Ala	Val	Gly	Cys	Gly	Pro	Trp
					395					400					405
5	Ser	Gln	Pro	Leu	Val	Val	Ser	Ser	His	Asp	Arg	Ala	Gly	Gln	Gln
					410					415					420
	Gly	Pro	Pro	His	Ser	Arg	Thr	Ser	Trp	Val	Pro	Val	Val	Leu	Gly
					425					430					435
10	Val	Leu	Thr	Ala	Leu	Val	Thr	Ala	Ala	Ala	Leu	Ala	Leu	Ile	Leu
					440					445					450
	Leu	Arg	Lys	Arg	Arg	Lys	Glu	Thr	Arg	Phe	Gly	Gln	Ala	Phe	Asp
					455					460					465
15	Ser	Val	Met	Ala	Arg	Gly	Glu	Pro	Ala	Val	His	Phe	Arg	Ala	Ala
					470					475					480
	Arg	Ser	Phe	Asn	Arg	Glu	Arg	Pro	Glu	Arg	Ile	Glu	Ala	Thr	Leu
					485					490					495
	Asp	Ser	Leu	Gly	Ile	Ser	Asp	Glu	Leu	Lys	Glu	Lys	Leu	Glu	Asp
					500					505					510
	Val	Leu	Ile	Pro	Glu	Gln	Gln	Phe	Thr	Leu	Gly	Arg	Met	Leu	Gly
					515					520					525
	Lys	Gly	Glu	Phe	Gly	Ser	Val	Arg	Glu	Ala	Gln	Leu	Lys	Gln	Glu
					530					535					540
	Asp	Gly	Ser	Phe	Val	Lys	Val	Ala	Val	Lys	Met	Leu	Lys	Ala	Asp
					545					550					555
	Ile	Ile	Ala	Ser	Ser	Asp	Ile	Glu	Glu	Phe	Leu	Arg	Glu	Ala	Ala
					560					565					570
	Cys	Met	Lys	Glu	Phe	Asp	His	Pro	His	Val	Ala	Lys	Leu	Val	Gly
					575					580					585
40	Val	Ser	Leu	Arg	Ser	Arg	Ala	Lys	Gly	Arg	Leu	Pro	Ile	Pro	Met
					590					595					600
	Val	Ile	Leu	Pro	Phe	Met	Lys	His	Gly	Asp	Leu	His	Ala	Phe	Leu
					605					610					615
45	Leu	Ala	Ser	Arg	Ile	Gly	Glu	Asn	Pro	Phe	Asn	Leu	Pro	Leu	Gln
					620					625					630
	Thr	Leu	Ile	Arg	Phe	Met	Val	Asp	Ile	Ala	Cys	Gly	Met	Glu	Tyr
					635					640					645

[illegible]

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3637 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

GAATTCTCGA GTCGACGTTG GACTTGAAGG AATGCCAAGA GATGCTGCCC 50
CCACCCCCTT AGGCCCGAGG GATCAGGAGC TATGGGACCA GAGGCCCTGT 100
CATCTTTACT GCTGCTGCTC TTGGTGGCAA GTGGAGATGC TGACATGAAG 150
GGACATTTTG ATCCTGCCAA GTGCCGCTAT GCCCTGGGCA TGCAGGACCG 200
GACCATCCCA GACAGTGACA TCTCTGCTTC CAGCTCCTGG TCAGATTCCA 250
CTGCCGCCCCG CCACAGCAGG TTGGAGAGCA GTGACGGGGA TGGGGCCTGG 300
TGCCCCGCAG GGTTCGGTGTT TCCCAAGGAG GAGGAGTACT TGCAGGTGGA 350
TCTACAACGA CTGCACCTGG TGGCTCTGGT GGGCACCCAG GGACGGCATG 400
CCGGGGGCCT GGGCAAGGAG TTCTCCCGGA GCTACCGGCT GCGTTACTCC 450
CGGGATGGTC GCCGCTGGAT GGGCTGGAAG GACCGCTGGG GTCAGGAGGT 500
GATCTCAGGC AATGAGGACC CTGAGGGAGT GGTGCTGAAG GACCTTGGGC 550
CCCCCATGGT TGCCCGACTG GTTCGCTTCT ACCCCCGGGC TGACCGGGTC 600
ATGAGCGTCT GTCTGCGGGT AGAGCTCTAT GGCTGCCTCT GGAGGGATGG 650
ACTCCTGTCT TACACCGCCC CTGTGGGGCA GACAATGTAT TTATCTGAGG 700

CCGTGTACCT CAACGACTCC ACCTATGACG GACATACCGT GGGCGGACTG 750

CAGTATGGGG GTCTGGGCCA GCTGGCAGAT GGTGTGGTGG GGCTGGATGA 800

CTTTAGGAAG AGTCAGGAGC TGCGGGTCTG GCCAGGCTAT GACTATGTGG 850

GATGGAGCAA CCACAGCTTC TCCAGTGGCT ATGTGGAGAT GGAGTTTGAG 900

TTTGACCGGC TGAGGGCCTT CCAGGCTATG CAGGTCCACT GTAACAACAT 950

GCACACGCTG GGAGCCCGTC TGCCTGGCGG GGTGGAATGT CGCTTCCGGC 1000

GTGGCCCTGC CATGGCCTGG GAGGGGGAGC CCATGCGCCA CAACCTAGGG 1050

GGCAACCTGG GGGACCCAG AGCCCGGGCT GTCTCAGTGC CCCTTGGCGG 1100

CCGTGTGGCT CGCTTTCTGC AGTGCCGCTT CCTCTTTGCG GGGCCCTGGT 1150

TACTCTTCAG CGAAATCTCC TTCATCTCTG ATGTGGTGAA CAATTCCTCT 1200

CCGGCACTGG GAGGCACCTT CCCGCCAGCC CCCTGGTGGC CGCCTGGCCC 1250

ACCTCCCACC AACTTCAGCA GCTTGGAGCT GGAGCCCAGA GGCCAGCAGC 1300

CCGTGGCCAA GCCCGAGGGG AGCCCGACCG CCATCCTCAT CGGCTGCCTG 1350

GTGGCCATCA TCCTGCTCCT GCTGCTCATC ATTGCCCTCA TGCTCTGGCG 1400

GCTGCACTGG CGCAGGCTCC TCAGCAAGGC TGAACGGAGG GTGTTGGAAG 1450

AGGAGCTGAC GGTTCACCTC TCTGTCCCTG GGGACACTAT CCTCATCAAC 1500

AACCGCCCAG GTCCTAGAGA GCCACCCCG TACCAGGAGC CCCGGCCTCG 1550

TGGGAATCCG CCCCACTCCG CTCCCTGTGT CCCCAATGGC TCTGCGTTGC 1600

01234567891011121314151617181920212223242526272829303132333435363738394041424344454647484950

TGCTCTCCAA TCCAGCCTAC CGCCTCCTTC TGGCCACTTA CGCCCGTCCC 1650

5 CCTCGAGGCC CGGGCCCCCC CACACCCGCC TGGGCCAAAC CCACCAACAC 1700

CCAGGCCTAC AGTGGGGACT ATATGGAGCC TGAGAAGCCA GGCGCCCCGC 1750

10 TTCTGCCCCC ACCTCCCCAG AACAGCGTCC CCCATTATGC CGAGGCTGAC 1800

ATTGTTACCC TGCAGGGCGT CACCGGGGGC AACACCTATG CTGTGCCTGC 1850

15 ACTGCCCCCA GGGGCAGTCG GGGATGGGCC CCCAGAGTG GATTTCCCTC 1900

20 GATCTCGACT CCGCTTCAAG GAGAAGCTTG GCGAGGGCCA GTTTGGGGAG 1950

GTGCACCTGT GTGAGGTCGA CAGCCCTCAA GATCTGGTCA GTCTTGATTT 2000

25 CCCCCTTAAT GTGCGTAAGG GACACCCTTT GCTGGTAGCT GTCAAGATCT 2050

TACGGCCAGA TGCCACCAAG AATGCCAGGA ATGATTCCT GAAAGAGGTG 2100

AAGATCATGT CGAGGCTCAA GGACCCAAAC ATCATTCGGC TGCTGGGCGT 2150

35 GTGTGTGCAG GACGACCCCC TCTGCATGAT TACTGACTAC ATGGAGAACG 2200

GCGACCTCAA CCAGTTCCTC AGTGCCCACC AGCTGGAGGA CAAGGCAGCC 2250

40 GAGGGGGCCC CTGGGGACGG GCAGGCTGCG CAGGGGCCCCA CCATCAGCTA 2300

CCCAATGCTG CTGCATGTGG CAGCCCAGAT CGCCTCCGGC ATGCGCTATC 2350

45 TGGCCACACT CAACTTTGTA CATCGGGACC TGGCCACGCG GAACTGCCTA 2400

50 GTTGGGGAAA ATTTACCAT CAAAATCGCA GACTTTGGCA TGAGCCGGAA 2450

CCTCTATGCT GGGGACTATT ACCGTGTGCA GGGCCGGGCA GTGCTGCCCCA 2500

TCCGCTGGAT GGCCTGGGAG TGCATCCTCA TGGGGAAGTT CACGACTGCG 2550

AGTGACGTGT GGGCCTTTGG TGTGACCCTG TGGGAGGTGC TGATGCTCTG 2600

TAGGGCCCAG CCCTTTGGGC AGCTCACCGA CGAGCAGGTC ATCGAGAACG 2650

CGGGGGAGTT CTTCCGGGAC CAGGGCCGGC AGGTGTACCT GTCCCGGCCG 2700

CCTGCCTGCC CGCAGGGCCT ATATGAGCTG ATGCTTCGGT GCTGGAGCCG 2750

GGAGTCTGAG CAGCGACCAC CCTTTTCCCA GCTGCATCGG TTCCTGGCAG 2800

AGGATGCACT CAACACGGTG TGAATCACAC ATCCAGCTGC CCCTCCCTCA 2850

GGGAGTGATC CAGGGGAAGC CAGTGACACT AAAACAAGAG GACACAATGG 2900

CACCTCTGCC CTTCCCCTCC CGACAGCCCA TCACCTCTAA TAGAGGCAGT 2950

GAGACTGCAG AAGCCCCTGT CGCCCACCCA GCTGGTCCTG TGGATGGGAT 3000

CCTCTCCACC CTCCTCTAGC CATCCCTTGG GGAAGGGTGG GGAGAAATAT 3050

AGGATAGACA CTGGACATGG CCCATTGGAG CACCTGGGCC CCACTGGACA 3100

ACACTGATTC CTGGAGAGGT GGCTGCGCCC CCAGCTTCTC TCTCCCTGTC 3150

ACACACTGGA CCCCACTGGC TGAGAATCTG GGGGTGAGGA GGACAAGAAG 3200

GAGAGGAAAA TGTTTCCTTG TGCCTGCTCC TGTACTTGTC CTCAGCTTGG 3250

GCTTCTTCCT CCTCCATCAC CTGAAACACT GGACCTGGGG GTAGCCCCGC 3300

CCCAGCCCTC AGTCACCCCC ACTTCCCACC TGCAGTCTTG TAGCTAGAAC 3350

TTCTCTAAGC CTATACGTTT CTGTGGAGTA AATATTGGGA TTGGGGGGAA 3400

5 AGAGGGAGCA ACGGCCCATATA GCCTTGGGGT TGGACATCTC TAGTGTAGCT 3450

GCCACATTGA TTTTCTATA ATCACTTGGG GTTTGTACAT TTTTGGGGGG 3500

10 AGAGACACAG ATTTTACAC TAATATATGG ACCTAGCTTG AGGCAATTTT 3550

15 AATCCCCTGC ACTAGGCAGG TAATAATAAA GGTGAGTTT TCCACAAAAA 3600

AAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAA 3637

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 913 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met Gly Pro Glu Ala Leu Ser Ser Leu Leu Leu Leu Leu Val
1 5 10 15
Ala Ser Gly Asp Ala Asp Met Lys Gly His Phe Asp Pro Ala Lys
20 25 30
Cys Arg Tyr Ala Leu Gly Met Gln Asp Arg Thr Ile Pro Asp Ser
35 40 45
Asp Ile Ser Ala Ser Ser Ser Trp Ser Asp Ser Thr Ala Ala Arg
50 55 60
His Ser Arg Leu Glu Ser Ser Asp Gly Asp Gly Ala Trp Cys Pro
65 70 75
Ala Gly Ser Val Phe Pro Lys Glu Glu Glu Tyr Leu Gln Val Asp
80 85 90
Leu Gln Arg Leu His Leu Val Ala Leu Val Gly Thr Gln Gly Arg
95 100 105
His Ala Gly Gly Leu Gly Lys Glu Phe Ser Arg Ser Tyr Arg Leu
110 115 120

0123456789
20
25
30
35

	Arg	Tyr	Ser	Arg	Asp	Gly	Arg	Arg	Trp	Met	Gly	Trp	Lys	Asp	Arg	
					125					130					135	
5	Trp	Gly	Gln	Glu	Val	Ile	Ser	Gly	Asn	Glu	Asp	Pro	Glu	Gly	Val	
					140					145					150	
	Val	Leu	Lys	Asp	Leu	Gly	Pro	Pro	Met	Val	Ala	Arg	Leu	Val	Arg	
					155					160					165	
10	Phe	Tyr	Pro	Arg	Ala	Asp	Arg	Val	Met	Ser	Val	Cys	Leu	Arg	Val	
					170					175					180	
	Glu	Leu	Tyr	Gly	Cys	Leu	Trp	Arg	Asp	Gly	Leu	Leu	Ser	Tyr	Thr	
					185					190					195	
15	Ala	Pro	Val	Gly	Gln	Thr	Met	Tyr	Leu	Ser	Glu	Ala	Val	Tyr	Leu	
					200					205					210	
	Asn	Asp	Ser	Thr	Tyr	Asp	Gly	His	Thr	Val	Gly	Gly	Leu	Gln	Tyr	
					215					220					225	
	Gly	Gly	Leu	Gly	Gln	Leu	Ala	Asp	Gly	Val	Val	Gly	Leu	Asp	Asp	
					230					235					240	
	Phe	Arg	Lys	Ser	Gln	Glu	Leu	Arg	Val	Trp	Pro	Gly	Tyr	Asp	Tyr	
					245					250					255	
	Val	Gly	Trp	Ser	Asn	His	Ser	Phe	Ser	Ser	Gly	Tyr	Val	Glu	Met	
					260					265					270	
	Glu	Phe	Glu	Phe	Asp	Arg	Leu	Arg	Ala	Phe	Gln	Ala	Met	Gln	Val	
					275					280					285	
	His	Cys	Asn	Asn	Met	His	Thr	Leu	Gly	Ala	Arg	Leu	Pro	Gly	Gly	
					290					295					300	
	Val	Glu	Cys	Arg	Phe	Arg	Arg	Gly	Pro	Ala	Met	Ala	Trp	Glu	Gly	
					305					310					315	
40	Glu	Pro	Met	Arg	His	Asn	Leu	Gly	Gly	Asn	Leu	Gly	Asp	Pro	Arg	
					320					325					330	
	Ala	Arg	Ala	Val	Ser	Val	Pro	Leu	Gly	Gly	Arg	Val	Ala	Arg	Phe	
					335					340					345	
45	Leu	Gln	Cys	Arg	Phe	Leu	Phe	Ala	Gly	Pro	Trp	Leu	Leu	Phe	Ser	
					350					355					360	
	Glu	Ile	Ser	Phe	Ile	Ser	Asp	Val	Val	Asn	Asn	Ser	Ser	Pro	Ala	
50					365					370					375	

20
25
30
35

	Leu Gly Gly Thr	Phe 380	Pro Pro Ala Pro	Trp 385	Trp Pro Pro Gly	Pro 390
5	Pro Pro Thr Asn	Phe 395	Ser Ser Leu Glu	Leu 400	Glu Pro Arg Gly	Gln 405
	Gln Pro Val Ala	Lys 410	Pro Glu Gly Ser	Pro 415	Thr Ala Ile Leu	Ile 420
10	Gly Cys Leu Val	Ala 425	Ile Ile Leu Leu	Leu 430	Leu Leu Ile Ile	Ala 435
	Leu Met Leu Trp	Arg 440	Leu His Trp Arg	Arg 445	Leu Leu Ser Lys	Ala 450
15	Glu Arg Arg Val	Leu 455	Glu Glu Glu Leu	Thr 460	Val His Leu Ser	Val 465
	Pro Gly Asp Thr	Ile 470	Leu Ile Asn Asn	Arg 475	Pro Gly Pro Arg	Glu 480
20	Pro Pro Pro Tyr	Gln 485	Glu Pro Arg Pro	Arg 490	Gly Asn Pro Pro	His 495
25	Ser Ala Pro Cys	Val 500	Pro Asn Gly Ser	Ala 505	Leu Leu Leu Ser	Asn 510
	Pro Ala Tyr Arg	Leu 515	Leu Leu Ala Thr	Tyr 520	Ala Arg Pro Pro	Arg 525
30	Gly Pro Gly Pro	Pro 530	Thr Pro Ala Trp	Ala 535	Lys Pro Thr Asn	Thr 540
35	Gln Ala Tyr Ser	Gly 545	Asp Tyr Met Glu	Pro 550	Glu Lys Pro Gly	Ala 555
	Pro Leu Leu Pro	Pro 560	Pro Pro Gln Asn	Ser 565	Val Pro His Tyr	Ala 570
40	Glu Ala Asp Ile	Val 575	Thr Leu Gln Gly	Val 580	Thr Gly Gly Asn	Thr 585
	Tyr Ala Val Pro	Ala 590	Leu Pro Pro Gly	Ala 595	Val Gly Asp Gly	Pro 600
45	Pro Arg Val Asp	Phe 605	Pro Arg Ser Arg	Leu 610	Arg Phe Lys Glu	Lys 615
50	Leu Gly Glu Gly	Gln 620	Phe Gly Glu Val	His 625	Leu Cys Glu Val	Asp 630

Arg Cys Trp Ser Arg Glu Ser Glu Gln Arg Pro Pro Phe Ser Gln
890 895 900

Leu His Arg Phe Leu Ala Glu Asp Ala Leu Asn Thr Val
905 910 913

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1164 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

GCAGGTCTGA AGCTCATGGG AGCCCCGGTG AAGCTGACAG TGTCTCAGGG 50
GCAGCCGGTG AAGCTCAACT GCAGTGTGGA GGGGATGGAG GAGCCTGACA 100
TCCAGTGGGT GAAGGATGGG GCTGTGGTCC AGAACTTGA CCAGTTGTAC 150
ATCCCAGTCA GCGAGCAGCA CTGGATCGGC TTCCTCAGCC TGAAGTCAGT 200
GGAGCGCTCT GACGCCGGCC GGTACTGGTG CCAGGTGGAG GATGGGGGTG 250
AAACCGAGAT CTCCCAGCCA GTGTGGCTCA CGGTAGAAGG TGTGCCATTT 300
TTCACAGTGG AGCCAAAAGA TCTGGCAGTG CCACCCAATG CCCCTTTCCA 350
ACTGTCTTGT GAGGCTGTGG GTCCCCCTGA ACCTGTTACC ATTGTCTGGT 400
GGAGAGGAAC TACGAAGATC GGGGGACCCG CTCCCTCTCC ATCTGTTTTA 450
AATGTAACAG GGGTGACCCA GAGCACCATG TTTTCCTGTG AAGCTCACAA 500
CCTAAAAGGC CTGGCCTCTT CTCGCACAGC CACTGTTTAC CTTCAAGCAC 550
TGCCTGCAGC CCCCTTCAAC ATCACCGTGA CAAAGCTTTC CAGCAGCAAC 600

GCTAGTGTGG CCTGGATGCC AGGTGCTGAT GGCCGAGCTC TGCTACAGTC 650

CTGTACAGTT CAGGTGACAC AGGCCCCAGG AGGCTGGGAA GTCCTGGCTG 700

TTGTGGTCCC TGTGCCCCCC TTTACCTGCC TGCTCCGGGA CCTGGTGCCT 750

GCCACCAACT ACAGCCTCAG GGTGCGCTGT GCCAATGCCT TGGGGCCCTC 800

TCCCTATGCT GACTGGGTGC CCTTTCAGAC CAAGGGTCTA GCCCCAGCCA 850

GCGCTCCCCA AAACCTCCAT GCCATCCGCA CAGATTCAGG CCTCATCTTG 900

GAGTGGGAAG AAGTGATCCC CGAGGCCCTT TTGGAAGGCC CCCTGGGACC 950

CTACAAACTG TCCTGGGTTC AAGACAATGG AACCCAGGAT GAGCTGACAG 1000

TGGAGGGGAC CAGGGCCAAT TTGACAGGCT GGGATCCCCA AAAGGACCTG 1050

ATCGTACGTG TGTGCGTCTC CAATGCAGTT GGCTGTGGAC CCTGGAGTCA 1100

GCCACTGGTG GTCTCTTCTC ATGACCGTGC AGGCCAGCAG GGCCCTCCTC 1150

ACAGCCGCAC ATCC 1164

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 388 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Ala	Gly	Leu	Lys	Leu	Met	Gly	Ala	Pro	Val	Lys	Leu	Thr	Val	Ser
1				5					10					15

Gln	Gly	Gln	Pro	Val	Lys	Leu	Asn	Cys	Ser	Val	Glu	Gly	Met	Glu
				20					25					30

0123456789 10 11 12 13 14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35

	Glu	Pro	Asp	Ile	Gln	Trp	Val	Lys	Asp	Gly	Ala	Val	Val	Gln	Asn	
					35					40					45	
5	Leu	Asp	Gln	Leu	Tyr	Ile	Pro	Val	Ser	Glu	Gln	His	Trp	Ile	Gly	
					50					55					60	
	Phe	Leu	Ser	Leu	Lys	Ser	Val	Glu	Arg	Ser	Asp	Ala	Gly	Arg	Tyr	
					65					70					75	
10	Trp	Cys	Gln	Val	Glu	Asp	Gly	Gly	Glu	Thr	Glu	Ile	Ser	Gln	Pro	
					80					85					90	
	Val	Trp	Leu	Thr	Val	Glu	Gly	Val	Pro	Phe	Phe	Thr	Val	Glu	Pro	
					95					100					105	
15	Lys	Asp	Leu	Ala	Val	Pro	Pro	Asn	Ala	Pro	Phe	Gln	Leu	Ser	Cys	
					110					115					120	
	Glu	Ala	Val	Gly	Pro	Pro	Glu	Pro	Val	Thr	Ile	Val	Trp	Trp	Arg	
20					125					130					135	
	Gly	Thr	Thr	Lys	Ile	Gly	Gly	Pro	Ala	Pro	Ser	Pro	Ser	Val	Leu	
					140					145					150	
25	Asn	Val	Thr	Gly	Val	Thr	Gln	Ser	Thr	Met	Phe	Ser	Cys	Glu	Ala	
					155					160					165	
	His	Asn	Leu	Lys	Gly	Leu	Ala	Ser	Ser	Arg	Thr	Ala	Thr	Val	His	
					170					175					180	
	Leu	Gln	Ala	Leu	Pro	Ala	Ala	Pro	Phe	Asn	Ile	Thr	Val	Thr	Lys	
					185					190					195	
30	Leu	Ser	Ser	Ser	Asn	Ala	Ser	Val	Ala	Trp	Met	Pro	Gly	Ala	Asp	
35					200					205					210	
	Gly	Arg	Ala	Leu	Leu	Gln	Ser	Cys	Thr	Val	Gln	Val	Thr	Gln	Ala	
					215					220					225	
40	Pro	Gly	Gly	Trp	Glu	Val	Leu	Ala	Val	Val	Val	Pro	Val	Pro	Pro	
					230					235					240	
	Phe	Thr	Cys	Leu	Leu	Arg	Asp	Leu	Val	Pro	Ala	Thr	Asn	Tyr	Ser	
					245					250					255	
45	Leu	Arg	Val	Arg	Cys	Ala	Asn	Ala	Leu	Gly	Pro	Ser	Pro	Tyr	Ala	
					260					265					270	
	Asp	Trp	Val	Pro	Phe	Gln	Thr	Lys	Gly	Leu	Ala	Pro	Ala	Ser	Ala	
50					275					280					285	

Pro Gln Asn Leu His Ala Ile Arg Thr Asp Ser Gly Leu Ile Leu
290 295 300

5 Glu Trp Glu Glu Val Ile Pro Glu Ala Pro Leu Glu Gly Pro Leu
305 310 315

Gly Pro Tyr Lys Leu Ser Trp Val Gln Asp Asn Gly Thr Gln Asp
320 325 330

10 Glu Leu Thr Val Glu Gly Thr Arg Ala Asn Leu Thr Gly Trp Asp
335 340 345

Pro Gln Lys Asp Leu Ile Val Arg Val Cys Val Ser Asn Ala Val
350 355 360

15 Gly Cys Gly Pro Trp Ser Gln Pro Leu Val Val Ser Ser His Asp
365 370 375

Arg Ala Gly Gln Gln Gly Pro Pro His Ser Arg Thr Ser
380 385 388

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1197 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GATGCTGACA TGAAGGGACA TTTTGATCCT GCCAAGTGCC GCTATGCCCT 50

GGGCATGCAG GACCGGACCA TCCCAGACAG TGACATCTCT GCTTCCAGCT 100

CCTGGTCAGA TTCCACTGCC GCCCGCCACA GCAGGTTGGA GAGCAGTGAC 150

GGGGATGGGG CCTGGTGCCC CGCAGGGTCG GTGTTTCCCA AGGAGGAGGA 200

GTACTTGCAG GTGGATCTAC AACGACTGCA CCTGGTGGCT CTGGTGGGCA 250

CCCAGGGACG GCATGCCGGG GGCCTGGGCA AGGAGTTCTC CCGGAGCTAC 300

50 CGGCTGCGTT ACTCCCGGGA TGGTCGCCGC TGGATGGGCT GGAAGGACCG 350

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 399 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

10 Asp Ala Asp Met Lys Gly His Phe Asp Pro Ala Lys Cys Arg Tyr
1 5 10 15
Ala Leu Gly Met Gln Asp Arg Thr Ile Pro Asp Ser Asp Ile Ser
20 25 30
15 Ala Ser Ser Ser Trp Ser Asp Ser Thr Ala Ala Arg His Ser Arg
35 40 45
Leu Glu Ser Ser Asp Gly Asp Gly Ala Trp Cys Pro Ala Gly Ser
50 55 60
Val Phe Pro Lys Glu Glu Glu Tyr Leu Gln Val Asp Leu Gln Arg
65 70 75
Leu His Leu Val Ala Leu Val Gly Thr Gln Gly Arg His Ala Gly
80 85 90
Gly Leu Gly Lys Glu Phe Ser Arg Ser Tyr Arg Leu Arg Tyr Ser
95 100 105
Arg Asp Gly Arg Arg Trp Met Gly Trp Lys Asp Arg Trp Gly Gln
110 115 120
Glu Val Ile Ser Gly Asn Glu Asp Pro Glu Gly Val Val Leu Lys
125 130 135
Asp Leu Gly Pro Pro Met Val Ala Arg Leu Val Arg Phe Tyr Pro
140 145 150
40 Arg Ala Asp Arg Val Met Ser Val Cys Leu Arg Val Glu Leu Tyr
155 160 165
Gly Cys Leu Trp Arg Asp Gly Leu Leu Ser Tyr Thr Ala Pro Val
170 175 180
45 Gly Gln Thr Met Tyr Leu Ser Glu Ala Val Tyr Leu Asn Asp Ser
185 190 195
Thr Tyr Asp Gly His Thr Val Gly Gly Leu Gln Tyr Gly Gly Leu
200 205 210
50

	Gly	Gln	Leu	Ala	Asp	Gly	Val	Val	Gly	Leu	Asp	Asp	Phe	Arg	Lys
					215					220					225
5	Ser	Gln	Glu	Leu	Arg	Val	Trp	Pro	Gly	Tyr	Asp	Tyr	Val	Gly	Trp
					230					235					240
	Ser	Asn	His	Ser	Phe	Ser	Ser	Gly	Tyr	Val	Glu	Met	Glu	Phe	Glu
					245					250					255
10	Phe	Asp	Arg	Leu	Arg	Ala	Phe	Gln	Ala	Met	Gln	Val	His	Cys	Asn
					260					265					270
	Asn	Met	His	Thr	Leu	Gly	Ala	Arg	Leu	Pro	Gly	Gly	Val	Glu	Cys
					275					280					285
15	Arg	Phe	Arg	Arg	Gly	Pro	Ala	Met	Ala	Trp	Glu	Gly	Glu	Pro	Met
					290					295					300
	Arg	His	Asn	Leu	Gly	Gly	Asn	Leu	Gly	Asp	Pro	Arg	Ala	Arg	Ala
					305					310					315
	Val	Ser	Val	Pro	Leu	Gly	Gly	Arg	Val	Ala	Arg	Phe	Leu	Gln	Cys
					320					325					330
	Arg	Phe	Leu	Phe	Ala	Gly	Pro	Trp	Leu	Leu	Phe	Ser	Glu	Ile	Ser
					335					340					345
	Phe	Ile	Ser	Asp	Val	Val	Asn	Asn	Ser	Ser	Pro	Ala	Leu	Gly	Gly
					350					355					360
	Thr	Phe	Pro	Pro	Ala	Pro	Trp	Trp	Pro	Pro	Gly	Pro	Pro	Pro	Thr
					365					370					375
	Asn	Phe	Ser	Ser	Leu	Glu	Leu	Glu	Pro	Arg	Gly	Gln	Gln	Pro	Val
					380					385					390
	Ala	Lys	Pro	Glu	Gly	Ser	Pro	Thr	Ala						
					395				399						

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3785 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CCTCCGCCAC CCTCCTCTCA GCGCTCGCGG GCCGGGCCCCG GCATGGTGCG 50

65520 " 65520 "

CTCCCTACGG CGACTGGGTG CCCTTTCAGA CAAAGGGCCT AGCGCCAGCC 1000
5 AGAGCTCCTC AGAATTTCCA TGCCATTTCGT ACCGACTCAG GCCTTATCCT 1050
GGAATGGGAA GAAGTGATTC CTGAAGACCC TGGGGAAGGC CCCCTAGGAC 1100
10 CTTATAAGCT GTCCTGGGTC CAAGAAAATG GAACCCAGGA TGAGCTGATG 1150
GTGGAAGGGA CCAGGGCCAA TCTGACCGAC TGGGATCCCC AGAAGGACCT 1200
15 GATTTTGCCT GTGTGTGCCT CCAATGCAAT TGGTGATGGG CCCTGGAGTC 1250
AGCCACTGGT GGTGTCTTCT CATGACCATG CAGGGAGGCA GGGCCCTCCC 1300
CACAGCCGCA CATCCTGGGT GCCTGTGGTC CTGGGCGTGC TCACCGCCCT 1350
20 GATCACAGCT GCTGCCTTGG CCCTCATCCT GCTTCGGAAG AGACGCAAGG 1400
AGACGCGTTT CGGGCAAGCC TTTGACAGTG TCATGGCCCG AGGGGAGCCA 1450
GCTGTACACT TCCGGGCAGC CCGATCTTTC AATCGAGAAA GGCCTGAACG 1500
25 CATTGAGGCC ACATTGGATA GCCTGGGCAT CAGCGATGAA TTGAAGGAAA 1550
AGCTGGAGGA TGTCTCATT CCAGAGCAGC AGTTCACCCT CGGTCCGATG 1600
40 TTGGGCAAAG GAGAGTTTGG ATCAGTGCGG GAAGCCCAGC TAAAGCAGGA 1650
AGATGGCTCC TTCGTGAAAG TGGCAGTGAA GATGCTGAAA GCTGACATCA 1700
45 TTGCCTCAAG CGACATAGAA GAGTTCCTCC GGGAAGCAGC TTGCATGAAG 1750
GAGTTTGACC ATCCACACGT GGCCAAGCTT GTTGGGGTGA GCCTCCGGAG 1800
50

CAGGGCTAAA GGTCGTCTCC CCATTCCCAT GGTCATCCTG CCCTTCATGA 1850

AACATGGAGA CTTGCACGCC TTTCTGCTCG CCTCCCGAAT CGGGGAGAAC 1900

CCTTTTAACC TGCCCCCTGCA GACCCTGGTC CGGTTTCATGG TGGACATTGC 1950

CTGTGGCATG GAGTACCTGA GCTCCCGGAA CTTTCATCCAC CGAGACCTAG 2000

CAGCTCGGAA TTGCATGCTG GCCGAGGACA TGACAGTGTG TGTGGCTGAT 2050

TTTGGACTCT CTCGGAATA CTATAGCGGG GACTATTATC GTCAGGGCTG 2100

TGCCTCCAAA TTGCCCCGTCA AGTGGCTGGC CCTGGAGAGC TTGGCTGACA 2150

ACTTGTATAC TGTACACAGT GATGTGTGGG CCTTCGGGGT GACCATGTGG 2200

GAGATCATGA CTCGTGGGCA GACGCCATAT GCTGGCATTG AAAATGCTGA 2250

GATTTACAAC TACCTCATCG GCGGGAACCG CCTGAAGCAG CCTCCGGAGT 2300

GCATGGAGGA AGTGTATGAT CTCATGTACC AGTGCTGGAG CGCCGACCCC 2350

AAGCAGCGCC CAAGCTTCAC GTGTCTGCGA ATGGAAGTGG AGAACATTCT 2400

GGGCCACCTG TCTGTGCTGT CCACCAGCCA GGACCCCTTG TACATCAACA 2450

TTGAGAGAGC TGAGCAGCCT ACTGAGAGTG GCAGCCCTGA GCTGCACTGT 2500

GGAGAGCGAT CCAGCAGCGA GGCAGGGGAC GGCAGTGGCG TGGGGGCAGT 2550

AGGTGGCATC CCCAGTGA CTCTGGTACAT CTTTCAGCCCC GGAGGGCTAT 2600

CCGAGTCACC AGGGCAGCTG GAGCAGCAGC CAGAAAGCCC CCTCAATGAG 2650

AACCAGAGGC TGTTGTTGCT GCAGCAAGGG CTAAGTGCCTC ACAGTAGCTG 2700

20
25
30
35
40
45
50

TTAACCTCA GGCAGAGGAA AGTTGGGGCC CCTGGCTCTG CTGACCGCTG 2750

5 CGCTGCCTGA CTAGGCCAG TCTGATCACA GCCCAGGCAG CAAGGTATGG 2800

AGGCTCCTGT GGTAGCCCTC CCAAGCTGTG TGGCGCCTGG ACGGACCAA 2850

10 TTGCCCAATC CCAGTTCTTC CTGCAGCCGC TCTGGCCAGC CTGGCATCAG 2900

TTCAGGCCTT GGCTTAGAGG AGGTGAGCCA GAGCTGGTTG CCTGAATGCA 2950

15 GGCAGCTGGC AGGAGGGGAG GGTGGCTATG TTTCCATGGG TACCATGGGT 3000

GTGGATGGCA GTAAGGGAGG GTAGCAACAG CCTGTGGGCC CCTACCCTCC 3050

TGGCTGAGCT GCTCCTACTT TAGTGCATGC TTGGAGCCGC CTGCAGCCTG 3100

GAACTCAGCA CTGCCCCACCA CACTTGGGCC GAAATGCCAG GTTTGCCCCT 3150

CTTAAGTCAC AAAGAGATGT CCATGTATTG TTCCCTTTTA GGTGATGATT 3200

AGGAAGGGAT TGGCACACTT GGGTCCCTAA GCCCTATGGC AGGAAATGGT 3250

35 GGGATATTCT CAGGTCTGAA TCCTCATCAT CTCCTGATT CCCCACCCTG 3300

CAAAGGCCTG GAACTGGCTG TGGGGCTCTG ACGCATGCTG AAGGACAAA 3350

40 GGTACAGAG ATCCGACTTC AAAAGGCAGG GTCTGAGTCT GGCAGGTGGA 3400

GAGGTGCTAA GGGGCTGGCC CAGGAGTCAG GCATTTTCAGG ACCCCTCCAA 3450

45 GCTTCTACAG TCTGTCTGAG CATGCTACCA AGCCCCAGA TACCCCAAAA 3500

50 CTAACAGAGG CAGTTTTGTC TGAGCCCAGC CCTCCACAT GATGACCCTT 3550

AGGTCTACCC TCCTCTCTAA ATGGACATCC TCGTTTGTCC CAAGTCTCCA 3600

GAGAGACTAC TGATGGCTGA TGTGGGTAAG AAAAGTTCCA GGAACCAGGG 3650

CTGGGGTGGA ACCAGGGCTG GGGTCGAGGC AGGCTCTTGG GCAGGCTCTT 3700

GCTGTTAGGA ACATTTCTAA GCTATTAAGT TGCTGTTTCA AAACAAATAA 3750

AATTGAAACA TAAAGAATCA AAAAAAAAAA AAAAA 3785

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 880 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Met	Ala	Leu	Arg	Arg	Ser	Met	Gly	Trp	Pro	Gly	Leu	Arg	Pro	Leu
1				5					10					15
Leu	Leu	Ala	Gly	Leu	Ala	Ser	Leu	Leu	Leu	Pro	Gly	Ser	Ala	Ala
			20						25					30
Ala	Gly	Leu	Lys	Leu	Met	Gly	Ala	Pro	Val	Lys	Met	Thr	Val	Ser
			35						40					45
Gln	Gly	Gln	Pro	Val	Lys	Leu	Asn	Cys	Ser	Val	Glu	Gly	Met	Glu
			50						55					60
Asp	Pro	Asp	Ile	His	Trp	Met	Lys	Asp	Gly	Thr	Val	Val	Gln	Asn
			65						70					75
Ala	Ser	Gln	Val	Ser	Ile	Ser	Ile	Ser	Glu	His	Ser	Trp	Ile	Gly
			80						85					90
Leu	Leu	Ser	Leu	Lys	Ser	Val	Glu	Arg	Ser	Asp	Ala	Gly	Leu	Tyr
			95						100					105
Trp	Cys	Gln	Val	Lys	Asp	Gly	Glu	Glu	Thr	Lys	Ile	Ser	Gln	Ser
			110						115					120
Val	Trp	Leu	Thr	Val	Glu	Gly	Val	Pro	Phe	Phe	Thr	Val	Glu	Pro
			125						130					135

	Lys	Asp	Leu	Ala	Val	Pro	Pro	Asn	Ala	Pro	Phe	Gln	Leu	Ser	Cys
					140					145					150
5	Glu	Ala	Val	Gly	Pro	Pro	Glu	Pro	Val	Thr	Ile	Tyr	Trp	Trp	Arg
					155					160					165
	Gly	Leu	Thr	Lys	Val	Gly	Gly	Pro	Ala	Pro	Ser	Pro	Ser	Val	Leu
					170					175					180
10	Asn	Val	Thr	Gly	Val	Thr	Gln	Arg	Thr	Glu	Phe	Ser	Cys	Glu	Ala
					185					190					195
	Arg	Asn	Ile	Lys	Gly	Leu	Ala	Thr	Ser	Arg	Pro	Ala	Ile	Val	Arg
					200					205					210
15	Leu	Gln	Ala	Pro	Pro	Ala	Ala	Pro	Phe	Asn	Thr	Thr	Val	Thr	Thr
					215					220					225
20	Ile	Ser	Ser	Tyr	Asn	Ala	Ser	Val	Ala	Trp	Val	Pro	Gly	Ala	Asp
					230					235					240
	Gly	Leu	Ala	Leu	Leu	His	Ser	Cys	Thr	Val	Gln	Val	Ala	His	Ala
					245					250					255
25	Pro	Gly	Glu	Trp	Glu	Ala	Leu	Ala	Val	Val	Val	Pro	Val	Pro	Pro
					260					265					270
	Phe	Thr	Cys	Leu	Leu	Arg	Asn	Leu	Ala	Pro	Ala	Thr	Asn	Tyr	Ser
					275					280					285
30	Leu	Arg	Val	Arg	Cys	Ala	Asn	Ala	Leu	Gly	Pro	Ser	Pro	Tyr	Gly
					290					295					300
	Asp	Trp	Val	Pro	Phe	Gln	Thr	Lys	Gly	Leu	Ala	Pro	Ala	Arg	Ala
35					305					310					315
	Pro	Gln	Asn	Phe	His	Ala	Ile	Arg	Thr	Asp	Ser	Gly	Leu	Ile	Leu
					320					325					330
40	Glu	Trp	Glu	Glu	Val	Ile	Pro	Glu	Asp	Pro	Gly	Glu	Gly	Pro	Leu
					335					340					345
	Gly	Pro	Tyr	Lys	Leu	Ser	Trp	Val	Gln	Glu	Asn	Gly	Thr	Gln	Asp
					350					355					360
45	Glu	Leu	Met	Val	Glu	Gly	Thr	Arg	Ala	Asn	Leu	Thr	Asp	Trp	Asp
					365					370					375
	Pro	Gln	Lys	Asp	Leu	Ile	Leu	Arg	Val	Cys	Ala	Ser	Asn	Ala	Ile
50					380					385					390

	Gly	Asp	Gly	Pro	Trp	Ser	Gln	Pro	Leu	Val	Val	Ser	Ser	His	Asp	
					395					400					405	
5	His	Ala	Gly	Arg	Gln	Gly	Pro	Pro	His	Ser	Arg	Thr	Ser	Trp	Val	
					410					415					420	
	Pro	Val	Val	Leu	Gly	Val	Leu	Thr	Ala	Leu	Ile	Thr	Ala	Ala	Ala	
					425					430					435	
10	Leu	Ala	Leu	Ile	Leu	Leu	Arg	Lys	Arg	Arg	Lys	Glu	Thr	Arg	Phe	
					440					445					450	
	Gly	Gln	Ala	Phe	Asp	Ser	Val	Met	Ala	Arg	Gly	Glu	Pro	Ala	Val	
					455					460					465	
15	His	Phe	Arg	Ala	Ala	Arg	Ser	Phe	Asn	Arg	Glu	Arg	Pro	Glu	Arg	
					470					475					480	
	Ile	Glu	Ala	Thr	Leu	Asp	Ser	Leu	Gly	Ile	Ser	Asp	Glu	Leu	Lys	
					485					490					495	
20	Glu	Lys	Leu	Glu	Asp	Val	Leu	Ile	Pro	Glu	Gln	Gln	Phe	Thr	Leu	
					500					505					510	
25	Gly	Arg	Met	Leu	Gly	Lys	Gly	Glu	Phe	Gly	Ser	Val	Arg	Glu	Ala	
					515					520					525	
	Gln	Leu	Lys	Gln	Glu	Asp	Gly	Ser	Phe	Val	Lys	Val	Ala	Val	Lys	
					530					535					540	
	Met	Leu	Lys	Ala	Asp	Ile	Ile	Ala	Ser	Ser	Asp	Ile	Glu	Glu	Phe	
					545					550					555	
	Leu	Arg	Glu	Ala	Ala	Cys	Met	Lys	Glu	Phe	Asp	His	Pro	His	Val	
					560					565					570	
	Ala	Lys	Leu	Val	Gly	Val	Ser	Leu	Arg	Ser	Arg	Ala	Lys	Gly	Arg	
					575					580					585	
40	Leu	Pro	Ile	Pro	Met	Val	Ile	Leu	Pro	Phe	Met	Lys	His	Gly	Asp	
					590					595					600	
	Leu	His	Ala	Phe	Leu	Leu	Ala	Ser	Arg	Ile	Gly	Glu	Asn	Pro	Phe	
					605					610					615	
45	Asn	Leu	Pro	Leu	Gln	Thr	Leu	Val	Arg	Phe	Met	Val	Asp	Ile	Ala	
					620					625					630	
	Cys	Gly	Met	Glu	Tyr	Leu	Ser	Ser	Arg	Asn	Phe	Ile	His	Arg	Asp	
					635					640					645	
50																

	Leu	Ala	Ala	Arg	Asn	Cys	Met	Leu	Ala	Glu	Asp	Met	Thr	Val	Cys
					650					655					660
5	Val	Ala	Asp	Phe	Gly	Leu	Ser	Arg	Lys	Ile	Tyr	Ser	Gly	Asp	Tyr
					665					670					675
	Tyr	Arg	Gln	Gly	Cys	Ala	Ser	Lys	Leu	Pro	Val	Lys	Trp	Leu	Ala
					680					685					690
10	Leu	Glu	Ser	Leu	Ala	Asp	Asn	Leu	Tyr	Thr	Val	His	Ser	Asp	Val
					695					700					705
	Trp	Ala	Phe	Gly	Val	Thr	Met	Trp	Glu	Ile	Met	Thr	Arg	Gly	Gln
					710					715					720
15	Thr	Pro	Tyr	Ala	Gly	Ile	Glu	Asn	Ala	Glu	Ile	Tyr	Asn	Tyr	Leu
					725					730					735
	Ile	Gly	Gly	Asn	Arg	Leu	Lys	Gln	Pro	Pro	Glu	Cys	Met	Glu	Glu
					740					745					750
20	Val	Tyr	Asp	Leu	Met	Tyr	Gln	Cys	Trp	Ser	Ala	Asp	Pro	Lys	Gln
					755					760					765
25	Arg	Pro	Ser	Phe	Thr	Cys	Leu	Arg	Met	Glu	Leu	Glu	Asn	Ile	Leu
					770					775					780
	Gly	His	Leu	Ser	Val	Leu	Ser	Thr	Ser	Gln	Asp	Pro	Leu	Tyr	Ile
					785					790					795
30	Asn	Ile	Glu	Arg	Ala	Glu	Gln	Pro	Thr	Glu	Ser	Gly	Ser	Pro	Glu
					800					805					810
	Leu	His	Cys	Gly	Glu	Arg	Ser	Ser	Ser	Glu	Ala	Gly	Asp	Gly	Ser
					815					820					825
	Gly	Val	Gly	Ala	Val	Gly	Gly	Ile	Pro	Ser	Asp	Ser	Arg	Tyr	Ile
					830					835					840
40	Phe	Ser	Pro	Gly	Gly	Leu	Ser	Glu	Ser	Pro	Gly	Gln	Leu	Glu	Gln
					845					850					855
	Gln	Pro	Glu	Ser	Pro	Leu	Asn	Glu	Asn	Gln	Arg	Leu	Leu	Leu	Leu
					860					865					870
45	Gln	Gln	Gly	Leu	Leu	Pro	His	Ser	Ser	Cys					
					875					880					

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
 (A) LENGTH: 1164 bases

(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

5 (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

GCAGGCCTGA AGCTCATGGG CGCCCCAGTG AAGATGACCG TGTCTCAGGG 50
GCAGCCAGTG AAGCTCAACT GCAGCGTGGA GGGGATGGAG GACCCTGACA 100
TCCACTGGAT GAAGGATGGC ACCGTGGTCC AGAATGCAAG CCAGGTGTCC 150
ATCTCCATCA GCGAGCACAG CTGGATTGGC TTACTCAGCC TAAAGTCAGT 200
GGAGCGGTCT GATGCTGGCC TGTACTGGTG CCAGGTGAAG GATGGGGAGG 250
AAACCAAGAT CTCTCAGTCA GTATGGCTCA CTGTCGAAGG TGTGCCATTC 300
TTCACAGTGG AACCAAAAGA TCTGGCGGTG CCACCCAATG CCCCTTTTCA 350
GCTGTCTTGT GAGGCTGTGG GTCCTCCAGA ACCCGTAACC ATTTACTGGT 400
GGAGAGGACT CACTAAGGTT GGGGGACCTG CTCCCTCTCC CTCTGTTTTA 450
AATGTGACAG GAGTGACCCA GCGCACAGAG TTTTCTTGTG AAGCCCGCAA 500
CATAAAAGGC CTGGCCACTT CCCGACCAGC CATTGTTCGC CTTCAAGCAC 550
CGCCTGCAGC TCCTTTCAAC ACCACAGTAA CAACGATCTC CAGCTACAAC 600
GCTAGCGTGG CCTGGGTGCC AGGTGCTGAC GGCCTAGCTC TGCTGCATTC 650
CTGTACTGTA CAGGTGGCAC ACGCCCCAGG AGAATGGGAG GCCCTTGCTG 700
TTGTGGTTCC TGTGCCACCT TTTACCTGCC TGCTTCGGAA CTTGGCCCCT 750

GCCACCAACT ACAGCCTTAG GGTGCGCTGT GCCAATGCCT TGGGCCCTTC 800
 TCCCTACGGC GACTGGGTGC CCTTTCAGAC AAAGGGCCTA GCGCCAGCCA 850
 GAGCTCCTCA GAATTTCCAT GCCATTCGTA CCGACTCAGG CCTTATCCTG 900
 GAATGGGAAG AAGTGATTCC TGAAGACCCT GGGGAAGGCC CCCTAGGACC 950
 TTATAAGCTG TCCTGGGTCC AAGAAAATGG AACCCAGGAT GAGCTGATGG 1000
 TGGAAGGGAC CAGGGCCAAT CTGACCGACT GGGATCCCCA GAAGGACCTG 1050
 ATTTTGCGTG TGTGTGCCTC CAATGCAATT GGTGATGGGC CCTGGAGTCA 1100
 GCCACTGGTG GTGTCTTCTC ATGACCATGC AGGGAGGCAG GGCCCTCCCC 1150
 ACAGCCGCAC ATCC 1164

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 388 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Ala	Gly	Leu	Lys	Leu	Met	Gly	Ala	Pro	Val	Lys	Met	Thr	Val	Ser
1				5					10					15
Gln	Gly	Gln	Pro	Val	Lys	Leu	Asn	Cys	Ser	Val	Glu	Gly	Met	Glu
				20					25					30
Asp	Pro	Asp	Ile	His	Trp	Met	Lys	Asp	Gly	Thr	Val	Val	Gln	Asn
				35					40					45
Ala	Ser	Gln	Val	Ser	Ile	Ser	Ile	Ser	Glu	His	Ser	Trp	Ile	Gly
				50					55					60
Leu	Leu	Ser	Leu	Lys	Ser	Val	Glu	Arg	Ser	Asp	Ala	Gly	Leu	Tyr
				65					70					75

Trp	Cys	Gln	Val	Lys 80	Asp	Gly	Glu	Glu	Thr 85	Lys	Ile	Ser	Gln	Ser 90
Val	Trp	Leu	Thr	Val 95	Glu	Gly	Val	Pro	Phe 100	Phe	Thr	Val	Glu	Pro 105
Lys	Asp	Leu	Ala	Val 110	Pro	Pro	Asn	Ala	Pro 115	Phe	Gln	Leu	Ser	Cys 120
Glu	Ala	Val	Gly	Pro 125	Pro	Glu	Pro	Val	Thr 130	Ile	Tyr	Trp	Trp	Arg 135
Gly	Leu	Thr	Lys	Val 140	Gly	Gly	Pro	Ala	Pro 145	Ser	Pro	Ser	Val	Leu 150
Asn	Val	Thr	Gly	Val 155	Thr	Gln	Arg	Thr	Glu 160	Phe	Ser	Cys	Glu	Ala 165
Arg	Asn	Ile	Lys	Gly 170	Leu	Ala	Thr	Ser	Arg 175	Pro	Ala	Ile	Val	Arg 180
Leu	Gln	Ala	Pro	Pro 185	Ala	Ala	Pro	Phe	Asn 190	Thr	Thr	Val	Thr	Thr 195
Ile	Ser	Ser	Tyr	Asn 200	Ala	Ser	Val	Ala	Trp 205	Val	Pro	Gly	Ala	Asp 210
Gly	Leu	Ala	Leu	Leu 215	His	Ser	Cys	Thr	Val 220	Gln	Val	Ala	His	Ala 225
Pro	Gly	Glu	Trp	Glu 230	Ala	Leu	Ala	Val	Val 235	Val	Pro	Val	Pro	Pro 240
Phe	Thr	Cys	Leu	Leu 245	Arg	Asn	Leu	Ala	Pro 250	Ala	Thr	Asn	Tyr	Ser 255
Leu	Arg	Val	Arg	Cys 260	Ala	Asn	Ala	Leu	Gly 265	Pro	Ser	Pro	Tyr	Gly 270
Asp	Trp	Val	Pro	Phe 275	Gln	Thr	Lys	Gly	Leu 280	Ala	Pro	Ala	Arg	Ala 285
Pro	Gln	Asn	Phe	His 290	Ala	Ile	Arg	Thr	Asp 295	Ser	Gly	Leu	Ile	Leu 300
Glu	Trp	Glu	Glu	Val 305	Ile	Pro	Glu	Asp	Pro 310	Gly	Glu	Gly	Pro	Leu 315
Gly	Pro	Tyr	Lys	Leu 320	Ser	Trp	Val	Gln	Glu 325	Asn	Gly	Thr	Gln	Asp 330

Glu Leu Met Val Glu Gly Thr Arg Ala Asn Leu Thr Asp Trp Asp
335 340 345

Pro Gln Lys Asp Leu Ile Leu Arg Val Cys Ala Ser Asn Ala Ile
350 355 360

Gly Asp Gly Pro Trp Ser Gln Pro Leu Val Val Ser Ser His Asp
365 370 375

His Ala Gly Arg Gln Gly Pro Pro His Ser Arg Thr Ser
380 385 388

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CGGATCCACA CGATGCGACT CTT 23

(2) INFORMATION FOR SEQ ID NO:14:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

GGAATTCCTC TCATGGAGCT AGTCCATCTC T 31

(2) INFORMATION FOR SEQ ID NO:15:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CGGATCCATC CACAGAGATG T 21

5 (2) INFORMATION FOR SEQ ID NO:16:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

15 GGAATTCCAA AGGACCAGCA CGATC 25

20 (2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

25 GACCGTGTGT GTGGCTGACT TTGGACTCTC CTGGAAGATC 40

35 (2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 40 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

45 GGCTGTGCCT CCAAATTGCC CGTCAAGTGG CTGGCCCTGG 40

50 (2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 50 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

AGCCGGTGAA GCTGAACTGC AGTGTGGAGG GGATGGAGGA GCCTGACATC 50

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 50 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

TCCAGCTACA ACGCTAGCGT GGCCTGGGTG CCAGGTGCTG ACGGCCTAGC 50

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 9 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Ile His Arg Asp Leu Ala Ala Arg Asn
1 5 9

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 6 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

Lys Trp Ile Ala Ile Glu
1 5 6

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 8 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

Lys Thr Trp Thr Met Ala Pro Glu
1 5 8

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Lys Trp Leu Ala Leu Glu
1 5 6

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 6 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Lys Trp Met Ala Leu Glu
1 5 6

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 30 bases

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

CAGCTGCTCG AGGCAGGTCT GAAGCTCATG 30

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GCATGAATTC ATGGCACACC TTCTACCGTG 30

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

CACTGAGCTG GCTGACTAAG 20

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CCTGATAGGC TGGGTACTCC 20

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

AAGCCCGGAC TGACCAAA 18

(2) INFORMATION FOR SEQ ID NO:31:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GTGCGGAATC AGAAAGATGG 20

(2) INFORMATION FOR SEQ ID NO:32:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 18 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

TCAAGACAAT GGAACCCA 18

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 36 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

CATGGAATTC GGTGACCGAT GTGCGGCTGT GAGGAG 36

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 894 amino acids

(B) TYPE: amino acid

(D) TOPOLOGY: linear

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

Met Ala Trp Arg Cys Pro Arg Met Gly Arg Val Pro Leu Ala Trp
1 5 10 15

Cys Leu Ala Leu Cys Gly Trp Ala Cys Met Ala Pro Arg Gly Thr
20 25 30

Gln Ala Glu Glu Ser Pro Phe Val Gly Asn Pro Gly Asn Ile Thr
35 40 45

Gly Ala Arg Gly Leu Thr Gly Thr Leu Arg Cys Gln Leu Gln Val
50 55 60

Gln Gly Glu Pro Pro Glu Val His Trp Leu Arg Asp Gly Gln Ile
65 70 75

Leu Glu Leu Ala Asp Ser Thr Gln Thr Gln Val Pro Leu Gly Glu
80 85 90

Asp Glu Gln Asp Asp Trp Ile Val Val Ser Gln Leu Arg Ile Thr
95 100 105

Ser Leu Gln Leu Ser Asp Thr Gly Gln Tyr Gln Cys Leu Val Phe
110 115 120

Leu Gly His Gln Thr Phe Val Ser Gln Pro Gly Tyr Val Gly Leu
125 130 135

Glu Gly Leu Pro Tyr Phe Leu Glu Glu Pro Glu Asp Arg Thr Val
140 145 150

Ala Ala Asn Thr Pro Phe Asn Leu Ser Cys Gln Ala Gln Gly Pro
155 160 165

Pro Glu Pro Val Asp Leu Leu Trp Leu Gln Asp Ala Val Pro Leu
170 175 180

Ala Thr Ala Pro Gly His Gly Pro Gln Arg Ser Leu His Val Pro
185 190 195

Gly Leu Asn Lys Thr Ser Ser Phe Ser Cys Glu Ala His Asn Ala
200 205 210

	Lys Gly Val Thr	Thr Ser Arg Thr Ala Thr Ile Thr Val Leu Pro	215	220	225
5	Gln Gln Pro Arg	Asn Leu His Leu Val Ser Arg Gln Pro Thr Glu	230	235	240
	Leu Glu Val Ala	Trp Thr Pro Gly Leu Ser Gly Ile Tyr Pro Leu	245	250	255
10	Thr His Cys Thr	Leu Gln Ala Val Leu Ser Asp Asp Gly Met Gly	260	265	270
	Ile Gln Ala Gly	Glu Pro Asp Pro Pro Glu Glu Pro Leu Thr Ser	275	280	285
15	Gln Ala Ser Val	Pro Pro His Gln Leu Arg Leu Gly Ser Leu His	290	295	300
	Pro His Thr Pro	Tyr His Ile Arg Val Ala Cys Thr Ser Ser Gln	305	310	315
	Gly Pro Ser Ser	Trp Thr His Trp Leu Pro Val Glu Thr Pro Glu	320	325	330
20	Gly Val Pro Leu	Gly Pro Pro Glu Asn Ile Ser Ala Thr Arg Asn	335	340	345
	Gly Ser Gln Ala	Phe Val His Trp Gln Glu Pro Arg Ala Pro Leu	350	355	360
25	Gln Gly Thr Leu	Leu Gly Tyr Arg Leu Ala Tyr Gln Gly Gln Asp	365	370	375
	Thr Pro Glu Val	Leu Met Asp Ile Gly Leu Arg Gln Glu Val Thr	380	385	390
30	Leu Glu Leu Gln	Gly Asp Gly Ser Val Ser Asn Leu Thr Val Cys	395	400	405
	Val Ala Ala Tyr	Thr Ala Ala Gly Asp Gly Pro Trp Ser Leu Pro	410	415	420
35	Val Pro Leu Glu	Ala Trp Arg Pro Gly Gln Ala Gln Pro Val His	425	430	435
	Gln Leu Val Lys	Glu Pro Ser Thr Pro Ala Phe Ser Trp Pro Trp	440	445	450
40	Trp Tyr Val Leu	Leu Gly Ala Val Val Ala Ala Ala Cys Val Leu	455	460	465

	Ile	Leu	Ala	Leu	Phe	Leu	Val	His	Arg	Arg	Lys	Lys	Glu	Thr	Arg	
					470					475					480	
5	Tyr	Gly	Glu	Val	Phe	Glu	Pro	Thr	Val	Glu	Arg	Gly	Glu	Leu	Val	
					485					490					495	
	Val	Arg	Tyr	Arg	Val	Arg	Lys	Ser	Tyr	Ser	Arg	Arg	Thr	Thr	Glu	
					500					505					510	
10	Ala	Thr	Leu	Asn	Ser	Leu	Gly	Ile	Ser	Glu	Glu	Leu	Lys	Glu	Lys	
					515					520					525	
	Leu	Arg	Asp	Val	Met	Val	Asp	Arg	His	Lys	Val	Ala	Leu	Gly	Lys	
					530					535					540	
15	Thr	Leu	Gly	Glu	Gly	Glu	Phe	Gly	Ala	Val	Met	Glu	Gly	Gln	Leu	
					545					550					555	
	Asn	Gln	Asp	Asp	Ser	Ile	Leu	Lys	Val	Ala	Val	Lys	Thr	Met	Lys	
					560					565					570	
	Ile	Ala	Ile	Cys	Thr	Arg	Ser	Glu	Leu	Glu	Asp	Phe	Leu	Ser	Glu	
					575					580					585	
25	Ala	Val	Cys	Met	Lys	Glu	Phe	Asp	His	Pro	Asn	Val	Met	Arg	Leu	
					590					595					600	
	Ile	Gly	Val	Cys	Phe	Gln	Gly	Ser	Glu	Arg	Glu	Ser	Phe	Pro	Ala	
					605					610					615	
	Pro	Val	Val	Ile	Leu	Pro	Phe	Met	Lys	His	Gly	Asp	Leu	His	Ser	
					620					625					630	
35	Phe	Leu	Leu	Tyr	Ser	Arg	Leu	Gly	Asp	Gln	Pro	Val	Tyr	Leu	Pro	
					635					640					645	
	Thr	Gln	Met	Leu	Val	Lys	Phe	Met	Ala	Asp	Ile	Ala	Ser	Gly	Met	
					650					655					660	
40	Glu	Tyr	Leu	Ser	Thr	Lys	Arg	Phe	Ile	His	Arg	Asp	Leu	Ala	Ala	
					665					670					675	
	Arg	Asn	Cys	Met	Leu	Asn	Glu	Asn	Met	Ser	Val	Cys	Val	Ala	Asp	
					680					685					690	
45	Phe	Gly	Leu	Ser	Lys	Lys	Ile	Tyr	Asn	Gly	Asp	Tyr	Tyr	Arg	Gln	
					695					700					705	
50	Gly	Arg	Ile	Ala	Lys	Met	Pro	Val	Lys	Trp	Ile	Ala	Ile	Glu	Ser	
					710					715					720	

Phe Val Gly Asn Pro Gly Asn Ile Thr Gly Ala Arg Gly Leu Thr
 35 40 45
 5 Gly Thr Leu Arg Cys Glu Leu Gln Val Gln Gly Glu Pro Pro Glu
 50 55 60
 Val Val Trp Leu Arg Asp Gly Gln Ile Leu Glu Leu Ala Asp Asn
 65 70 75
 10 Thr Gln Thr Gln Val Pro Leu Gly Glu Asp Trp Gln Asp Glu Trp
 80 85 90
 Lys Val Val Ser Gln Leu Arg Ile Ser Ala Leu Gln Leu Ser Asp
 95 100 105
 15 Ala Gly Glu Tyr Gln Cys Met Val His Leu Glu Gly Arg Thr Phe
 110 115 120
 20 Val Ser Gln Pro Gly Phe Val Gly Leu Glu Gly Leu Pro Tyr Phe
 125 130 135
 Leu Glu Glu Pro Glu Asp Lys Ala Val Pro Ala Asn Thr Pro Phe
 140 145 150
 25 Asn Leu Ser Cys Gln Ala Gln Gly Pro Pro Glu Pro Val Thr Leu
 155 160 165
 Leu Trp Leu Gln Asp Ala Val Pro Leu Ala Pro Val Thr Gly His
 170 175 180
 30 Ser Ser Gln His Ser Leu Gln Thr Pro Gly Leu Asn Lys Thr Ser
 185 190 195
 35 Ser Phe Ser Cys Glu Ala His Asn Ala Lys Gly Val Thr Thr Ser
 200 205 210
 Arg Thr Ala Thr Ile Thr Val Leu Pro Gln Arg Pro His His Leu
 215 220 225
 40 His Val Val Ser Arg Gln Pro Thr Glu Leu Glu Val Ala Trp Thr
 230 235 240
 Pro Gly Leu Ser Gly Ile Tyr Pro Leu Thr His Cys Asn Leu Gln
 245 250 255
 45 Ala Val Leu Ser Asp Asp Gly Val Gly Ile Trp Leu Gly Lys Ser
 260 265 270
 50 Asp Pro Pro Glu Asp Pro Leu Thr Leu Gln Val Ser Val Pro Pro
 275 280 285

5 His Gln Leu Arg Leu Glu Lys Leu Leu Pro His Thr Pro Tyr His
 290 295 300
 Ile Arg Ile Ser Cys Ser Ser Ser Gln Gly Pro Ser Pro Trp Thr
 305 310 315
 10 His Trp Leu Pro Val Glu Thr Thr Glu Gly Val Pro Leu Gly Pro
 320 325 330
 Pro Glu Asn Val Ser Ala Met Arg Asn Gly Ser Gln Val Leu Val
 335 340 345
 15 Arg Trp Gln Glu Pro Arg Val Pro Leu Gln Gly Thr Leu Leu Gly
 350 355 360
 Tyr Arg Leu Ala Tyr Arg Gly Gln Asp Thr Pro Glu Val Leu Met
 365 370 375
 20 Asp Ile Gly Leu Thr Arg Glu Val Thr Leu Glu Leu Arg Gly Asp
 380 385 390
 Arg Pro Val Ala Asn Leu Thr Val Ser Val Thr Ala Tyr Thr Ser
 395 400 405
 25 Ala Gly Asp Gly Pro Trp Ser Leu Pro Val Pro Leu Glu Pro Trp
 410 415 420
 Arg Pro Gly Gln Gly Gln Pro Leu His His Leu Val Ser Glu Pro
 425 430 435
 30 Pro Pro Arg Ala Phe Ser Trp Pro Trp Trp Tyr Val Leu Leu Gly
 440 445 450
 35 Ala Leu Val Ala Ala Ala Cys Val Leu Ile Leu Ala Leu Phe Leu
 455 460 465
 Val His Arg Arg Lys Lys Glu Thr Arg Tyr Gly Glu Val Phe Glu
 470 475 480
 40 Pro Thr Val Glu Arg Gly Glu Leu Val Val Arg Tyr Arg Val Arg
 485 490 495
 Lys Ser Tyr Ser Arg Arg Thr Thr Glu Ala Thr Leu Asn Ser Leu
 500 505 510
 45 Gly Ile Ser Glu Glu Leu Lys Glu Lys Leu Arg Asp Val Met Val
 515 520 525
 50 Asp Arg His Lys Val Ala Leu Gly Lys Thr Leu Gly Glu Gly Glu
 530 535 540

5 Phe Gly Ala Val Met Glu Gly Gln Leu Asn Gln Asp Asp Ser Ile
 545 550 555
 Leu Lys Val Ala Val Lys Thr Met Lys Ile Ala Ile Cys Thr Arg
 560 565 570
 10 Ser Glu Leu Glu Asp Phe Leu Ser Glu Ala Val Cys Met Lys Glu
 575 580 585
 Phe Asp His Pro Asn Val Met Arg Leu Ile Gly Val Cys Phe Gln
 590 595 600
 15 Gly Ser Asp Arg Glu Gly Phe Pro Glu Pro Val Val Ile Leu Pro
 605 610 615
 Phe Met Lys His Gly Asp Leu His Ser Phe Leu Leu Tyr Ser Arg
 620 625 630
 20 Leu Gly Asp Gln Pro Val Phe Leu Pro Thr Gln Met Leu Val Lys
 635 640 645
 Phe Met Ala Asp Ile Ala Ser Gly Met Glu Tyr Leu Ser Thr Lys
 650 655 660
 25 Arg Phe Ile His Arg Asp Leu Ala Ala Arg Asn Cys Met Leu Asn
 665 670 675
 Glu Asn Met Ser Val Cys Val Ala Asp Phe Gly Leu Ser Lys Lys
 680 685 690
 30 Ile Tyr Asn Gly Asp Tyr Tyr Arg Gln Gly Arg Ile Ala Lys Met
 695 700 705
 35 Pro Val Lys Trp Ile Ala Ile Glu Ser Leu Ala Asp Arg Val Tyr
 710 715 720
 Thr Ser Lys Ser Asp Val Trp Ser Phe Gly Val Thr Met Trp Glu
 725 730 735
 40 Ile Ala Thr Arg Gly Gln Thr Pro Tyr Pro Gly Val Glu Asn Ser
 740 745 750
 Glu Ile Tyr Asp Tyr Leu Arg Gln Gly Asn Arg Leu Lys Gln Pro
 755 760 765
 45 Val Asp Phe Leu Asp Gly Leu Tyr Ser Leu Met Ser Arg Cys Trp
 770 775 780
 50 Glu Leu Asn Pro Arg Asp Arg Pro Ser Phe Ala Glu Leu Arg Glu
 785 790 795

Asp Leu Glu Asn Thr Leu Lys Ala Leu Pro Pro Ala Gln Glu Pro
800 805 810

5 Asp Glu Ile Leu Tyr Val Asn Met Asp Glu Gly Gly Ser His Leu
815 820 825

Glu Pro Arg Gly Ala Ala Gly Gly Ala Asp Pro Pro Thr Gln Pro
830 835 840

10 Asp Pro Lys Asp Ser Cys Ser Cys Leu Thr Ala Ala Asp Val His
845 850 855

Ser Ala Gly Arg Tyr Val Leu Cys Pro Ser Thr Ala Pro Gly Pro
860 865 870

15 Thr Leu Ser Ala Asp Arg Gly Cys Pro Ala Pro Pro Gly Gln Glu
875 880 885

Asp Gly Ala
888

00236939-012599
655270-6E69E220